

CHAPITRE V - ANALYSE LOCALE DE DONNEES EN TERMES DE TESTS DE SIGNIFICATION :  
UN EXEMPLE ET COMMENTAIRES

Dans ce chapitre, nous procéderons à une première analyse inférentielle des "données H & B", dans la perspective d'analyse locale évoquée au chapitre précédent. Dans cette analyse, nous adopterons le critère d'évaluation (traditionnel) des tests de signification. (Nous présenterons au chapitre VII des analyses qui reprendront les mêmes questions mais selon un nouveau critère d'évaluation).

Nous supposerons donc, pour fixer les idées, que les effets sur lesquels on souhaite concentrer l'examen correspondent d'une part à la question A.B/C (avec  $C = C_2$  : c'est-à-dire la question principale de la recherche originale), d'autre part aux deux questions B/A\*c1 et B/A\*c2 (deux questions secondaires choisies ici à titre illustratif), appliquées d'une part au protocole de groupe fondamental  $S*A*B*C_2$ , d'autre part à chacun des protocoles individuels fondamentaux (du type  $R<A*B*C>/s1$ ) : (cf. tableau III), d'où finalement la planification des analyses locales suivante :

- { trois analyses locales sur le protocole de groupe fondamental ;
- { trois analyses locales sur chacun des protocoles individuels fondamentaux.

Dans ce chapitre, nous présenterons d'abord le détail des procédures pour les interrogations relatives au protocole de groupe, puis l'ensemble des résultats.

Auparavant, nous résumerons quelques idées de base relatives à l'inférence statistique en général et aux tests de signification en particulier (ceci pour rendre le texte aussi autonome que possible ; pour les lecteurs statisticiens, il s'agira, bien sûr, de rappels).

(1) Toute procédure inférentielle fait appel à un modèle d'échantillonnage, selon lequel les observations sont regardées comme provenant de populations "parentes" ; dans l'exemple présent :

- (pour les analyses individuelles) : population d'épreuves, ou répétitions ;
- (en outre pour les analyses de groupe) : population de sujets.

Dans cette perspective, chaque effet observé est regardé en quelque sorte comme le "reflet statistique" de l'effet "parent" correspondant, défini au niveau des populations.

(2) Quant au critère des tests de signification, il consiste à examiner l'hypothèse d'absence d'effet au niveau de la population, hypothèse dite classiquement hypothèse nulle.

Dans ce but, on considère une statistique de test, dont la distribution d'échantillonnage, sous l'hypothèse nulle, coïncide avec celle d'une "variable de référence" (telle que le "t de Student" ou le "F de Snedecor" disons, dans ce qui suit immédiatement, le t de Student ; selon l'usage, nous utiliserons la même lettre, en l'occurrence t, pour la statistique de test et la variable de référence). On calcule, à partir des données, la valeur de la statistique de test ; si cette valeur est proche d'une valeur centrale de la distribution de référence, on conclut que les données sont compatibles avec l'hypothèse nulle d'absence d'effet (parent) ; si au contraire elle en est éloignée, on conclut à l'existence d'un effet (parent), ou, comme on dit encore, on rejette l'hypothèse nulle d'absence d'effet.

Les conclusions au test sont traditionnellement agrémentées de précisions quantitatives fondées sur la notion de "statistique significative au seuil  $\alpha$ ". Désignons, toujours selon l'usage, par  $t_\alpha$  la valeur qui n'est dépassée (en valeur absolue) que par une proportion  $\alpha$  des valeurs de la distribution de référence (la valeur  $t_\alpha$  est appelée la "valeur critique de la distribution du t de Student au seuil  $\alpha$ ) ; on dit que la statistique de test est "significative au seuil  $\alpha$ " si elle est supérieure (en valeur absolue) à  $t_\alpha$ , ce qu'on écrit : " $p < \alpha$ ", la lettre p désignant ce qu'on appelle le "seuil observé". En vue d'une conclusion de rejet, on envisage couramment comme seuils  $\alpha$  les seuils suivants, que nous appellerons "seuils conventionnels" : 0.05, 0.01 ou 0.001 ; plus le résultat est significatif, plus la conclusion de rejet apparaîtra étayée par les données. A l'opposé, la conclusion de compatibilité est de même souvent agrémentée d'une précision indiquant que la statistique de test n'est "pas significative" (sous-entendu : aux seuils conventionnels, ou, mieux encore, à des seuils tels que 0.10, 0.20, etc.). En langue anglaise, les phrases suivantes ont le statut de formulations consacrées :

"There is  $\left\{ \begin{array}{l} \text{a strong} \\ \text{some} \\ \text{no} \end{array} \right\}$  évidence of an effect :  $\left\{ \begin{array}{l} p < 0.001 \\ p < 0.05 \\ p > 0.20 \end{array} \right\}$  "

Remarque : Le lecteur déjà familiarisé avec la statistique trouvera certainement très élémentaires les procédures mises en oeuvre ci-dessous ; en effet : (1) nous envisageons seulement des "effets à 1 degré de liberté",

d'où la possibilité d'utiliser la variable de référence plus simple "t de Student" au lieu du "F de Snedecor" ; (2) nous nous bornons à envisager les termes adjoints minimaux : ceux qui dans une théorie plus générale du test F (cf. Réf. 1970 et M. Waisbrot, 1977) conduisent aux rapports notés (dans les sorties du programme VAR3)  $F'$  ou  $F_1$ . Mais ici, notre objectif principal n'est pas d'exposer des techniques mais d'illustrer une démarche d'analyse ; or, dans des situations plus complexes (plusieurs degrés de liberté, non-unicité des termes adjoints), cette démarche restera fondamentalement la même, même si la complexité accrue entraîne des points de choix nouveaux, que permettra précisément d'aborder la formalisation linéaire (cf. chapitre 6).

Analyses locales sur le protocole de groupe fondamental S\*A\*B\*C

1) Test de signification de l'effet A.B/C (avec  $C = C_2$ )

La question A.B/C, appliquée au protocole de groupe fondamental, S\*A\*B\*C, conduit à l'interrogation A.B/C\*S (cf. Tableau III). L'analyse en termes de test de signification de cette interrogation comportera les étapes suivantes :

- pour chaque session, on calcule pour chaque sujet l'effet d'interaction entre les facteurs A et B, puis on prend la moyenne des effets sur les sessions. Ainsi, pour le sujet s1, les effets d'interaction sont :

pour la session c1 :  $(357.7 - 388.7) - (372.0 - 406.7) = 3.7$   
pour la session c2 :  $(326.3 - 347.8) - (371.2 - 394.3) = 1.5 ;$

d'où l'effet moyenné sur les deux sessions (on prend  $C = E_2$ ) :  
 $1/2 (3.7 + 1.5) = 2.6.$

Le protocole (dérivé) des effets d'interaction individuels moyennés sur les sessions est présenté dans le tableau suivant (valeurs arrondies ; rappelons que les sujets ont été numérotés selon l'ordre croissant des valeurs absolues des effets d'interaction, c'est-à-dire à partir de la considération de ce protocole dérivé).

s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	Moy.
3	6	7	9	11	-11	-22	-24	25	-29	-29	-35	-7.4

Ce protocole a pour moyenne  $m = -7.4$  et pour écart-type-corrigé  $s = 19.8$ . Nous appellerons ce protocole : protocole dérivé pertinent pour l'interrogation A.B/C) et les statistiques  $m$  et  $s$  : statistiques pertinentes ; ce sont elles qui serviront de base à toutes les analyses inférentielles.

Comme statistique de test, on peut prendre ici  $t = \frac{m}{s/\sqrt{n}}$  (avec  $n-1$  d.l.,  $n$  étant le nombre de sujets, ici 12) ; on trouve  $t = \frac{-7.4}{19.8/\sqrt{12}} = -1.29$ . Cette valeur est non-significative (au seuil  $\alpha = 0.10$ , la valeur critique (bilatérale)  $t_{\alpha}$  vaut 1.80, d'où  $p > 0.10$ ) ; on conclut que les données sont compatibles avec l'hypothèse d'absence (dans la population parente des sujets) d'effet A.B/C (effet d'interaction entre les facteurs principaux A et B, moyenné sur les 2 sessions  $c1$  et  $c2$ ).

2) Test de signification de l'effet B/A\*c1

La question B/A\*c1, appliquée au protocole de groupe fondamental, conduit à l'interrogation B/A\*c1\*S ; le test de signification comportera les étapes suivantes :

(1) On se restreint au sous-protocole S\*A\*B/c1 (4 premières colonnes du "tableau des données" de la sortie n° 1 de l'Annexe) ;

(2) On moyenne sur le facteur A, d'où le protocole dérivé S\*B/A\*c1.

(N.B. : nous adoptons ici la solution du moyennage équipondéré).

Ce protocole comporte pour chaque sujet, et pour chacune des modalités  $b1$  et  $b2$ , la valeur obtenue en moyennant les 2 temps de réaction du sujet relatifs aux 2 modalités du facteur A, pour la session  $c1$  ; ainsi pour le sujet  $s1$  :

$$\begin{cases} 1/2 (357.7 + 388.7) = 373.2 \\ 1/2 (372.0 + 406.7) = 389.4 \end{cases}$$

Ci-dessous, le tableau du protocole dérivé S+B/A+c1, avec en marge les valeurs des différences (lesquelles représentent, pour chaque sujet, l'effet du facteur  $B_2$ ).

	s1	s2	s3	s4	s5	s6	s7	s8	s9	s10	s11	s12	Moy.
b1	373	400	421	329	362	411	348	346	341	387	377	448	378.5
b2	389	468	467	356	403	444	374	366	381	373	448	488	413.1

Différences -16 -68 -46 -27 -41 -33 -27 -20 -39 +15 -72 -41 -34.6

Le protocole des différences constituera, pour la question B/A\*c1, le protocole dérivé pertinent ; sa moyenne  $m = -34.6$  et son écart-type-corrigé  $s = 23.0$  seront ici les statistiques pertinentes. La statistique de test vaut ici  $t = \frac{-34.6}{23.0/\sqrt{12}} = -5.22$  ; la valeur de  $|t|$  est supérieure à la valeur critique 4.337 au seuil  $\alpha = 0.001$  ; la différence entre les deux moyennes est donc très significative ( $p < 0.001$ ) ; on conclut donc à l'existence (dans la population parente des sujets) d'un effet B/A\*c1 (effet du facteur B moyenné sur A, pour la session c1).

### 3) Test de signification de l'effet B/A\*c2

On trouverait de même  $t = 6.54$  (11 d.l.), très significatif ( $p < 0.001$ ) conclusion similaire.

### Analyses locales par le programme VAR3

Les analyses locales précédentes, effectuées par le programme VAR3 figurent en Annexe, dans la sortie n° 2 ("analyses locales planifiées").

(1) Comme formules d'interrogation, on a utilisé  $A.B/C$  et  $B/A*c1$  (donc la notation par défaut pour le facteur S ; dans le programme VAR3, la notation pleine ne serait pas acceptée, parce que S est un facteur de groupe).

Par contre, on aurait pu simplement demander  $A.B$  au lieu de  $A.B/C$  (ce qu'on peut vérifier sur la sortie, où la formule  $A.B$  figure plus loin dans l'"analyse standard") ; et également  $B/c1$  au lieu de  $B/A*c1$  ;

(2) Pour chaque demande, le programme calcule, dans le cas présent, 2 rapports F notés F' et F" ; le rapport F' (égal ici à 1.68) est le carré de la statistique t calculée précédemment (ce rapport est obtenu en utilisant la source adjointe minimale, et sera valide sous un modèle "local" : cf. ci-dessous, p.71 ainsi que la "brochure verte" pour des commentaires généraux sur la validité des tests F' et F")

Analyses locales pour un protocole individuel R<A\*B\*C>: tests de signifi-  
cation des effets

Pour un protocole individuel noté ici R<A\*B\*C> la démarche sera analogue ; cependant, les procédures seront différentes des précédentes, puisque cette fois le facteur de groupe est emboîté au lieu d'être croisé.

Nous commenterons ici les résultats individuels du sujet s1 ; on trouvera, dans le tableau IV, les résultats pour chacun des 12 sujets en ce qui concerne l'effet A.B/C.

. Effet A.B/C : on trouve  $t = 0.17$  (avec 371 d.l.), valeur proche de zéro ; les données sont donc compatibles avec l'hypothèse d'absence d'effet (dans la population parente des épreuves du sujet s1) ;

. Effets B/A\*c1 et B/A\*c2 : de même que pour les analyses descriptives effectuées au chapitre 4, on peut procéder à une analyse pondérée ou équipondérée. Pour l'effet B/A\*c1, on trouve respectivement (définition pondérée)  $t = -1.53$  et (définition équipondérée)  $t = -1.42$  (avec 184 d.l. dans les deux cas), valeurs non-significatives ( $p > 0.10$ ). Pour l'effet B/A\*c2, on trouve respectivement  $t = -5.04$  et  $t = -4.39$  (avec 187 d.l.), valeurs fortement significatives ( $p < 0.001$ ) : on conclura donc à l'existence d'un effet B/A\*c2.

{ Avec le programme VAR3, on trouverait ici, pour chacun des effets B/A\*c1  
{ et B/A\*c2 (et pour chacune des versions pondérée ou équipondérée); deux  
{ rapports notés  $F_1$  et  $F_2$  ; le rapport  $F_1$  est le carré de la statistique  $t$   
{ rapportée ici (l'interprétation des rapports  $F_1$  et  $F_2$  appellerait des com-  
{ mentaires similaires à ceux évoqués plus haut à propos de  $F'$  et  $F''$ ).

COMMENTAIRES ET DISCUSSION (\*)

L'ensemble des analyses locales que nous venons d'effectuer sur les "données H & B" (nous dirons brièvement : l'"analyse locale" de ces données) a apporté des réponses précises (moyennant le critère d'évaluation des tests de signification) à toutes les interrogations posées au début de ce chapitre ; elle pourra donc être considérée, vis-à-vis des objectifs fixés par la recherche, comme "complète" (du moins vis-à-vis du critère d'évaluation adopté).

---

(+) Cette partie (qui peut être omise sans inconvénient pour la lecture de la suite) suppose une certaine familiarité avec les pratiques ou théories "usuelles".

Dans ce paragraphe, nous situerons la perspective de l'"Analyse locale des données", des points de vue pratique et théorique, par rapport à ce que nous avons appelé au chapitre 1er "pratiques établies" et "théorie reçue". Pour illustrer la discussion, nous prendrons l'analyse du protocole de groupe fondamental.

Les pratiques établies conduiraient ici à procéder d'abord à l'analyse standard du protocole. Le plan a ici la structure statistique  $S \times T$ , où le facteur  $T$  est le croisement  $A \times B \times C$  ; l'analyse standard consiste à examiner, selon le critère des tests de signification, chacun des 7 effets liés à la décomposition canonique du croisement  $A \times B \times C$  (cf. chapitre IV et Annexe, sortie n° 2).

A chacun des 7 effets, on peut associer un effet adjoint minimal, qui sera ici l'interaction entre l'effet et le facteur de groupe  $S$ . Par ailleurs, l'effet adjoint maximal, commun à tous les effets, est l'interaction entre le facteur composé  $A \times B \times C$  et le facteur de groupe  $S$ .

L'analyse standard consiste à calculer, pour chacun des 7 effets : une somme de carrés (ou inertie), puis un carré moyen et un rapport  $F$  ; en ce qui concerne le dénominateur de ce dernier, on rencontre dans les pratiques usuelles les deux variantes suivantes :

- dans l'une, on prend le carré moyen associé à l'effet adjoint minimal correspondant à l'effet examiné, d'où le rapport que nous désignons ici par  $F'$  ;

- dans l'autre, on prend un dénominateur unique : le carré moyen associé à l'effet adjoint maximal, d'où le rapport que nous désignons ici par  $F''$ .

#### Commentaires relatifs au programme VAR3

(1) Pour tout plan ayant la structure statistique  $S \times C \times T$  on pourra toujours effectuer l'analyse standard, et même, avec la dernière version (novembre 1977) du programme, au moyen d'une demande unique qui engendrera automatiquement tous les termes de la décomposition canonique des effets liés au plan déclaré ;

(2) Le programme envisage plusieurs variantes pour les dénominateurs, qui dans l'exemple présent conduiront aux deux rapports  $F'$  et  $F''$  (cf. "brochure verte") pour précisions complémentaires).

Cependant, dans la perspective de l'analyse locale des données, l'analyse standard n'apparaîtra appropriée que lorsque les termes qui figurent dans la décomposition coïncident avec les interrogations que l'expérimentaliste souhaite poser aux données.

(1) Mais fréquemment vis-à-vis de ces dernières, l'analyse standard apparaîtra insuffisante. Ainsi, dans le cas présent : l'interrogation principale : A.B/C figure bien dans l'analyse standard, mais non les interrogations secondaires B/A\*c1 et B/A\*c2. D'une manière générale, pour que l'on puisse envisager toutes les questions au moyen de l'analyse standard, il faudrait que ces questions soient toutes formulées à un même niveau de restriction.

Dans certains cas, il sera possible de remplacer un ensemble de questions que l'on se pose par un ensemble de questions figurant dans l'analyse standard ; par exemple, ici, remplacer :

$$\begin{cases} B/A*c1 \\ B/A*c2 \end{cases} \text{ par } \begin{cases} B/A*C \\ B.C/A, \end{cases}$$

questions qui figurent dans l'analyse standard (à l'écriture près).

En effet (cf. chapitre IV), chacun des ensembles de questions précédentes constitue une décomposition de l'effet B(C)/A : effet de B intra-C moyenné sur A, c'est-à-dire ;

$$\begin{cases} B(c1)/A \\ B(c2)/A \end{cases} \text{ qui peut s'écrire } \begin{cases} B/A*c1 \\ B/A*c2 \end{cases} \text{ ou encore } \begin{cases} B/c1 \\ B/c2 \end{cases}$$

et

$$\begin{cases} B/C*A \\ B.C/A \end{cases} \text{ qui peut s'écrire } \begin{cases} B/A*C \\ B.C/A \end{cases} \text{ ou encore } \begin{cases} B \\ B.C \end{cases}$$

L'utilisateur du programme VAR3 pourra vérifier les synonymies des écritures précédentes, qui toutes constituent des demandes d'analyse admissibles par le programme (aux lettres minuscules près : cf. sorties de l'Annexe).

Dans la sortie n° 2 on a fait figurer (à la suite de l'analyse standard), la demande B(C)/A. Numériquement, on peut vérifier que l'inertie (somme des carrés) associée à B(C)/A (31293.4637) est égale :

- d'une part à la somme des deux inerties associées à B/A\*c1 et B/A\*c2 : 14368.9650 + 16924.4987 ;

- d'autre part, à la somme des deux inerties associées à B et B.C : 31241 + 52.2607 (cela, parce que le plan  $S \times A \times B \times C$  étant équilibré vis-à-vis de B et C, la décomposition en effet global et effet d'interaction est orthogonale).

Cependant, la "substitution" précédente n'apporte pas une solution véritable au problème, car si l'ensemble des 2 questions B/A\*C et B.C/A peut être regardé comme équivalent à l'ensemble des 2 questions initiales, il n'en demeure pas moins que chacune de ces deux questions initiales n'est pas envisagée dans le cadre de l'analyse standard. Encore le présent exemple est-il relativement simple, en ce sens qu'un utilisateur "chevronné" saurait sans doute "s'en tirer", en réinterprétant les résultats fournis par l'analyse standard dans les termes des questions initiales, à partir d'un raisonnement tel que le suivant : "puisque l'interaction (B.C) est non-significative, ( $F'=0.20$ ), c'est que les deux effets B/c1 et B/c2 doivent avoir à peu près le même degré de significativité ; et comme par ailleurs l'effet global (B) est très significatif ( $F'=47.30$ ) c'est qu'ils doivent être tous les deux très significatifs". Mais on voit tout le caractère approximatif, et partant mal généralisable, d'un tel raisonnement, qui atteint vite ses limites dès que le plan est tant soit peu complexe ; ainsi, la réinterprétation d'interactions doubles ou a fortiori d'ordre supérieur peut devenir à la fois très malaisée et périlleuse.

(2) L'analyse standard peut proposer, en revanche, des analyses locales superflues en ce sens qu'elles répondent à des questions que les objectifs de la recherche n'amèneraient guère à considérer. Ce dernier inconvénient pourrait, sans doute, être regardé comme mineur, si à ce moment, la "théorie reçue" ne venait pas interférer avec les pratiques. Nous développerons ce point plus loin lorsque nous aborderons le problème de la validité des méthodes inférentielles (et opposerons la notion de modèle général à celle de modèles locaux et dérivés).

(3) Les expérimentalistes, en général, sont conscients des limitations de l'analyse standard - d'où la pratique, mentionnée au chapitre I, consistant à juxtaposer, au tableau classique d'"analyse de la variance", des analyses complémentaires, qui, elles, visent expressément à répondre aux questions qui se posent effectivement. A ce point, les pratiques diffèrent selon le degré plus ou moins grand de "sophistication statistique" des utilisateurs.

Un utilisateur "peu sophistiqué" se livrera volontiers à des analyses "sauvages", c'est-à-dire effectuées sans prendre en compte la structure du plan, soit qu'il effectue de simples "comparaisons à vue" purement impressionnistes, soit (ce qui sans doute est pire parce que paré d'une garantie scientifique illusoire) qu'il utilise des procédures qui, valides pour des protocoles "élémentaires", ne le sont plus pour des protocoles complexes et peuvent alors donner lieu à des faussages incontrôlables (résultats indûment significatifs, etc.). Le prototype de ces "analyses sauvages" est la pratique des "comparaisons entre traitements deux à deux", par la méthode du "t de Student", pratique qui conduit à des résultats faussés, cela d'autant plus que le nombre de traitements est plus élevé.

Une autre pratique contestable est l'usage systématique des tests "non-paramétriques", supposés souvent (imprudemment) "libres de toute hypothèse", en réalité susceptibles de conduire également à toutes sortes de faussages.

Un utilisateur plus "sophistiqué" évitera les pièges précédents, en mettant en oeuvre des procédures inférentielles plus valides : par exemple, la méthode des comparaisons "a posteriori". En gros, procéder à une comparaison a posteriori revient à consulter, pour le même écart observé à une hypothèse nulle, de nouvelles tables beaucoup moins enclines que celles de la distribution du F à procurer des résultats significatifs ; moyennant quoi, on peut adopter la "règle" (avec toutes les implications ascétiques du terme) consistant à traiter toutes les comparaisons qui peuvent intéresser l'expérimentaliste mais qui n'entrent pas dans l'analyse standard (donc ici :  $A/B \times c_1$  et  $A/B \times c_2$ ) comme des comparaisons "a posteriori". Malheureusement, une telle règle, non seulement s'accorde aussi mal que possible avec la dynamique réelle de la démarche expérimentale, mais elle apparaît, dans nombre de situations, tout à fait arbitraire, voire incohérente (comme dans la recherche précédente, où nous avons supposé que l'examen des effets  $B/A \times c_1$  et  $B/A \times c_2$  avait bel et bien été envisagé "a priori"). Un tel caractère d'arbitraire apparaît inhérent à nombre de pratiques courantes, dont la mise en oeuvre provoque souvent de véritables "détournements d'intention" vis-à-vis des objectifs de la recherche expérimentale. Au lieu de fournir à l'expérimentaliste des méthodes qui l'amèneraient à concentrer son attention sur les questions à poser aux données elles-mêmes, ces pratiques renvoient constamment à des "problèmes" qui se présentent comme "orthogonaux" aux objectifs de la recherche et dont

la formulation est telle qu'il apparaît bien difficile d'y répondre, et pour cause. Par exemple : quel seuil de signification "aurait-on dû" choisir a priori pour savoir, lorsqu'on se trouve dans la situation "douteuse" d'un seuil observé  $p$  entre 0.05 et 0.10, si "on est en droit" ou non de conclure au rejet de l'hypothèse nulle ? Ou encore (toujours dans une situation douteuse) : n'aurait-on pas pu considérer que l'hypothèse "alternative" était orientée ? ("problème" dit du "choix" du test unilatéral ou bilatéral). Ou encore : quelles interactions pouvaient-elles être supposées nulles avant l'expérience ? (variante du "problème" dit du "pooling des dénominateurs", toujours posé dans l'espoir de grappiller un point ou deux supplémentaires de "significativité". Concluons : l'expérimentaliste cherchait des procédures objectives lui permettant de "dire l'information apportée par ses données", et il est constamment renvoyé à l'analyse de ses états de conscience antérieurs au recueil de ces données ...

Problèmes de validité ; modèle d'échantillonnage général et modèles locaux.

Toute méthode inférentielle soulève le problème de ses conditions de validité, c'est-à-dire les conditions qu'il convient de poser sur le modèle pour que les conclusions inférentielles soient correctes. Dans le cas des tests de signification, les conditions de validité spécifieront les conditions que doit vérifier le modèle d'échantillonnage pour que la statistique de test soit (sous l'hypothèse nulle) distribuée selon la distribution de référence.

[Une précision avant d'entreprendre la discussion ; les conditions de validité qu'on pose généralement sont, pour des raisons de commodité sur lesquelles nous reviendrons, des conditions "suffisantes" mais (il s'en faut souvent de beaucoup) non-nécessaires ; un point crucial de la discussion consistera précisément à montrer qu'il est à la fois possible et intéressant d'affaiblir dans une large mesure ces conditions].

Reprenons par exemple la procédure exposée à propos du protocole  $S \times A \times B \times C$ , pour examiner la question  $A.B/C$  ; à partir du protocole dérivé (numérique) des interactions individuelles, que nous désignerons ici par  $(z_s)_s \in S$ , nous avons défini la statistique de test  $t = \frac{m}{s/\sqrt{v}}$ . Pour que

cette statistique  $t$  soit distribuée selon la distribution de référence du  $t$  de Student, donc pour qu'elle définisse un test de signification valide, il suffit de poser le modèle d'échantillonnage suivant : les  $n$  observations numériques  $(z_s)_{s \in S}$  sont indépendantes et proviennent d'une distribution normale (i.e. gaussienne).

Un tel modèle d'échantillonnage, posé au niveau du protocole dérivé, constitue ce que nous appellerons un modèle local. "Modèle local" s'opposera à "modèle général" (c'est-à-dire posé au niveau du protocole fondamental) de même qu'"analyse locale" s'oppose à "analyse générale".

Dans le cas présent, le modèle général serait posé au niveau du protocole  $S \times A \times B \times C$ , ou  $S \times J$ , qu'on peut regarder comme constitué de  $n$  observations  $J$ -dimensionnelles ("vectorielles")  $(x^s)_{s \in S}$  avec  $x^s = (x^{sj})_{j \in J}$ , ("modèle mixte" selon la terminologie classique ; cf. Scheffé, chapitre VIII) ; et les conditions de validité seraient posées sur le modèle d'échantillonnage suivant : les  $n$  observations "vectorielles"  $(x^s)_{s \in S}$  sont indépendantes et proviennent d'une distribution multinormale (à  $/J/$  dimensions) ayant pour "vecteur" moyen  $(\mu^k)_{k \in K}$ . Du modèle général (posé sur le protocole  $(x^s)_{s \in S}$ , on déduirait un modèle pour le protocole dérivé  $(z^s)_{s \in S}$  : ce que nous appellerons le modèle dérivé relatif à ce protocole ; bien entendu, ce modèle dérivé coïncidera avec le modèle local posé directement plus haut.

D'une manière générale, toutes les fois que la théorie traditionnelle conduit à poser un modèle général, on aura l'implication ;

Modèle général  $\longrightarrow$  Modèle local

et le test de l'hypothèse nulle locale, valide sous le modèle local, sera donc également valide sous le modèle général.

Mais l'implication réciproque, manifestement, est fausse en général. D'où l'avantage, s'agissant d'une analyse inférentielle locale, de pouvoir spécifier les seules conditions de validité relatives au modèle local.

En règle générale, plus la question locale est simple (techniquement : moins elle comporte de degrés de liberté), moins le modèle local sera restrictif. Dans le cas d'un seul degré de liberté, la seule condition cruciale est celle de l'indépendance des observations (la condition de normalité, notamment, est très secondaire, comme l'ont amplement montré toutes les études statistiques effectuées sur le problème dit de la "robustesse des tests"). Il en résulte que le modèle local pourra se trouver admissible (et donc con-

duire à une inférence locale valide) dans des situations où le modèle général serait tout à fait inadmissible. En pareil cas, on voit que la préoccupation exclusive des conditions de validité du modèle général amènerait à jeter un doute souvent injustifié sur la validité de chacune des conclusions locales et en particulier sur celles qui correspondent aux objectifs de la recherche.

### Conditions de validité des analyses inférentielles locales dans la structure

#### $S\langle G\rangle * T$

Malgré l'importance pratique des structures statistiques  $S * T$  et  $S\langle G\rangle * T$ , ce n'est que récemment qu'ont été établies (cf. Réf. 1970, pour la structure  $S * T$  et M. Waisbrot, 1977, pour la structure  $S\langle G\rangle * T$ ), les conditions minimales à poser pour que les distributions d'échantillonnage des carrés moyens qu'on peut associer à une analyse locale soient distribuées comme des khi-deux, et conduisent donc à des tests F valides. On trouvera le détail de ces conditions dans les références précédentes (et un résumé dans la "brochure verte") ; mais ce qu'il convient de souligner ici, c'est que les conditions de validité à poser sur le modèle local diffèrent selon l'analyse locale envisagée. Au contraire, les conditions de validité du modèle général impliquent la validité de toutes les analyses locales. On comprend mieux, dès lors, où est l'"avantage" d'en rester, pour les considérations de validité, au niveau du modèle général ; cet "avantage" est ni plus ni moins celui de la simplicité, qui rendra commodes aussi bien des présentations théoriques de style général et introductif, que des procédures de calcul numérique simplifiées ; ce n'est certainement pas un hasard si les premiers textes d'analyse de la variance, antérieurs à l'"ère informatique", posaient par exemple, pour la structure statistique  $S * T$ , des conditions extrêmement strictes, généralement irréalistes mais conduisant à un seul et unique dénominateur pour tous les rapports F (en l'occurrence, le rapport que nous désignons par F").

#### Inférence locale et exhaustivité partielle

Comme nous l'avons dit, la perspective de l'analyse générale est celle qui est presque toujours mise en avant dans les développements théoriques traditionnels ("modèle général" posé d'emblée, etc.). Cela ne signifie pas, pour autant, que l'idée d'analyse locale soit totalement absente

de ces développements. Une fois dégagée, cette idée sera facilement "reconnue comme sous-jacente à maint exposé classique ; un exemple patent est la présentation usuelle du "t de Student pour groupes appareillés", où l'on pose toujours le modèle au niveau du protocole dérivé des différences (cf. par exemple la présentation de ce test, à propos des données de Student lui-même, dans le traité classique de Cramér (\*), p. 463).

D'une manière générale mais beaucoup plus difficile, ce qu'on pourrait appeler un "principe d'inférence locale" apparaîtra toujours comme pré-supposé, implicitement, chaque fois qu'une méthode d'inférence statistique est simplement envisagée comme applicable à des données. L'exemple suivant a clairement une valeur générale : supposons un sondage où on observe, sur un échantillon d'individus, les valeurs d'un ensemble K de variables (numériques pour fixer les idées), à partir desquelles on effectue des inférences diverses (estimations, tests, etc.) disons, sur les moyennes de la population parente.

Presque toujours, l'ensemble K des variables observées ne constitue qu'une partie d'un ensemble de variables "qu'on aurait pu observer". Donc nécessairement, en procédant à des inférences à partir des seules observations relatives aux variables K, on présuppose implicitement, que les conclusions ne seraient pas (ou pas trop) affectées si on disposait d'observations supplémentaires concernant des variables de l'ensemble  $\mathcal{X}-K$ . En d'autres termes, on présuppose que le protocole effectivement analysé, qui peut être regardé comme un sous-protocole du protocole relatif à toutes les variables de  $\mathcal{X}$ , est, vis-à-vis des paramètres relatifs à K, non seulement "pertinent" (contenant "de l'information" sur ces paramètres), mais "exhaustif" (contenant "toute l'information" sur ces paramètres). Nous dirons que lorsqu'une telle présupposition est faite, explicitement ou implicitement, on met en jeu, un "principe d'inférence locale".

A ce point de la discussion, certains pourraient trouver, du moins à première vue, que dans la plupart des situations un tel "principe" est non seulement raisonnable, mais va "presque de soi". Cependant, quelque réflexion conduira certainement à plus de prudence. Reprenons par exemple la situation précédente, et imaginons que certains paramètres relatifs à  $\mathcal{X}-K$  se trouvent liés fonctionnellement à des paramètres de K ; manifestement, avec cette contrainte, le recueil d'observations supplémentaires relatives aux variables correspondantes apporterait de l'information sur les paramètres relatifs à K ;

---

(\*) H. CRAMÉR (1961) *Mathematical methods of statistics*, Princeton, Univ. Pres.

en d'autres termes, le protocole relatif aux variables de K ne serait pas exhaustif vis-à-vis de ces paramètres, et le "principe d'inférence locale" ne serait pas applicable.

Du point de vue de l'inférence statistique, la problématique sous-jacente à l'analyse locale met clairement en jeu la notion que les statisticiens désignent sous le nom d'exhaustivité partielle (c'est-à-dire l'exhaustivité d'une statistique vis-à-vis d'un paramètre dérivé ; or, comme il est bien connu parmi les spécialistes, le cadre de l'inférence statistique "orthodoxe" se prête mal à la définition d'une "bonne notion" d'exhaustivité partielle ; il est à cet égard révélateur que ce soit seulement dans des textes de statistique bayésienne (où la formalisation de l'idée d'exhaustivité partielle est plus accessible), qu'on trouve des discussions relevant incontestablement de ce que nous appelons l'"analyse locale" (\*).

Dans nos propres travaux, on peut dire que c'est la place centrale que nous avons donnée à la notion d'analyse locale qui nous a conduits, profondément, à regarder comme inéluctable, l'élargissement du cadre "orthodoxe" de l'inférence statistique (Cf. chapitre VII, et Réf. 1977d).

#### QUELQUES PRINCIPES METHODOLOGIQUES POUR GUIDER L'ANALYSE INFERENTIELLE DES DONNEES EXPERIMENTALES

De la longue discussion qui précède, nous retiendrons quelques principes méthodologiques concernant la phase des "analyses fines" inférentielles de l'analyse des données expérimentales.

(1) A la juxtaposition d'une analyse standard "aveugle" et d'analyses complémentaires disparates, on préférera l'examen d'un ensemble unique d'analyses locales soigneusement réfléchies, découlant directement des objectifs de la recherche effectuée : ce que nous appelons un ensemble d'analyses locales planifiées.

(2) Lorsque plusieurs procédures sont envisageables pour examiner une

---

(\*) Cf. notamment (toujours à propos de la comparaison de deux groupes appariés) la discussion in D.V. LINDLEY (1965) : *Introduction to Probability and Statistics from a Bayesian Viewpoint*, Part 2, Inférence, Cambridge : Cambridge University Press.

question jugée essentielle, les considérations de validité (a-t-on "le droit" d'utiliser telle ou telle procédure ?) seront toujours subordonnées à celles de pertinence (a-t-on "intérêt" à utiliser telle ou telle procédure ?).

(3) Lorsque une (ou plusieurs) procédures envisageables apparaissent pertinentes, on pourra, en revanche, "jouer au plus fin" en recherchant notamment, sous quelles conditions, aussi peu restrictives que possible, ces procédures sont valides.

Il va de soi que la portée effective des principes précédents sera appréciée en tenant compte des particularités de chaque situation expérimentale ; mais en gros, on peut dire qu'un ensemble d'analyses locales planifiées se situera quelque part entre les deux cas extrêmes suivants :

- un ensemble de quelques "coups de sonde" donnés en plusieurs points précis, choisis pour leur importance toute particulière (on rejoindra ici la motivation à l'origine des "analyses complémentaires") ;

- un ensemble d'interrogations se rapprochant d'une "décomposition générale" (dans les cas où l'objectif essentiel du traitement serait plutôt de parvenir à une sorte de "tableau synoptique" des données : on rejoindra alors la motivation à l'origine de l'analyse standard). Mais bien entendu, dans ce cas, il n'y aura aucune raison a priori de privilégier la décomposition canonique, du moment que dans la planification des analyses, on peut tenir compte des éventuelles hiérarchies entre facteurs (vis-à-vis des objectifs de la recherche) pour parvenir à des résultats plus directement interprétables (cf. notamment ce qui a été dit aux chap. IV et V à propos de la décomposition d'un croisement).

Par ailleurs, dans les cas intermédiaires (les plus courants), rien n'empêchera le cas échéant d'adjoindre, aux demandes locales essentielles, d'autres demandes qui la "complèteront" de manière à reconstituer une analyse générale.

Par exemple, dans le cas des "données H & B", on pourrait envisager les analyses locales suivantes :

C  
  { A/B\*c1  
  { A/B\*c2  
  
  { B/A\*c1  
  { B/A\*c2  
  
  { A.B/C  
  { A.B.C

dont l'ensemble constitue une décomposition générale qui "complète" les 3 analyses locales relatives aux questions que nous avons posées comme questions essentielles.

Les considérations qui précèdent devraient, pensons-nous, éclairer un problème fréquemment soulevé par les expérimentalistes : celui du "nombre maximum" d'analyses inférentielles qu'il est recommandable d'effectuer.

[Ce problème est tout à fait général, mais il est parfois ressenti avec une certaine acuité par les utilisateurs de nos programmes, précisément parce que les possibilités offertes par ces programmes : langage de demandes et procédures d'analyse automatiques, se prêtent "trop bien" à la tentation de "facilité" consistant à "multiplier" les demandes d'analyse].

Pratiquement, disons d'abord que d'après notre propre expérience, il semble assez rare que le nombre de questions véritablement intéressantes dépasse de beaucoup, disons, celui qui correspondrait à une décomposition générale (l'un des deux "cas extrêmes" mentionnés plus haut). Or, l'idée de multiplier "sans raison" les demandes d'analyse va manifestement à l'encontre de la perspective (correctement entendue) de l'analyse locale inférentielle des données.

A fortiori, cette perspective implique qu'on ne multiplie pas les demandes d'analyse "pour de mauvaises raisons", c'est-à-dire, en clair, pour pratiquer une recherche systématique aveugle des résultats "significatifs". [Les chercheurs avertis, nous l'avons déjà signalé, savent, ne serait-ce que par expérience, que cette douteuse "pêche à la ligne" ne peut que conduire à des conclusions faussées, souvent grossièrement ; le fait-même que cette "pêche à la ligne" devienne avec les programmes de la série VAR, un "jeu d'enfant", devrait précisément, croyons-nous, mettre en alerte les chercheurs sur sa futilité].

Ce qu'il convient ici de reconnaître, c'est que le "problème des frontières entre les analyses "motivées" et les analyses "aveugles" est laissé à l'appréciation de l'utilisateur. Cependant, il ne faudrait pas voir là un point faible de la méthodologie, mais plutôt l'un des points sur lesquels celle-ci met l'utilisateur en face de ses véritables responsabilités ; alors que la "règle", mentionnée plus haut, stipulant de traiter comme "a posteriori" toute comparaison qui ne figurerait pas dans l'analyse standard ne laissait

d'où la possibilité de poser des questions précises quelle que soit par ailleurs la complexité du plan, la démarche d'interrogation restant toujours d'une grande simplicité conceptuelle. A quiconque a tenté d'analyser ou même simplement d'organiser des données tant soit peu complexes, l'intérêt, voire la nécessité pratique d'un tel langage apparaîtra clairement, même indépendamment de sa "valeur opératoire" qui ouvre la possibilité d'automatisation. En effet, il convient également de souligner les aspects sémantiques de ce langage : le rôle d'une formule, laquelle renvoie à la fois à une question (l'effet à examiner) et à une action (les dérivations à effectuer) ne se réduit pas à un simple intermédiaire de calcul (ce que sont notamment les "design matrices" des théories traditionnelles) ; et à cet égard, il est révélateur que les utilisateurs de nos programmes se comportent spontanément, vis-à-vis du langage des formules, comme vis-à-vis d'un langage "naturel" (avec les possibilités de nuances offertes par les synonymies, etc.).

Quant au privilège de la structure statistique  $S\langle G \rangle * T$ , il réside dans le fait que cette structure fournit d'une manière canonique, pour chaque effet défini par les facteurs systématiques, des "termes adjoints" qui pourront servir de référence pour l'évaluation de cet effet, et permettront notamment de guider les analyses inférentielles.

Du point de vue des programmes-machine, rappelons que les programmes d'exploration de données actuellement réalisés (programmes de V. Duquenne) mettent en jeu la seule structure de plan quasi-complet, et que de leur côté les programmes VAR3 et VAR4 sont fondés sur l'exploitation conjointe des deux structures privilégiées : le protocole traité doit être décrit par un plan muni de la structure statistique  $S\langle G \rangle * T$ , et les facteurs systématiques G et T doivent eux-mêmes, vis-à-vis des facteurs élémentaires (ce qu'on appellera les "facteurs déclarés du plan"), constituer une structure de facteur que complet (équilibré).

Quelques principes généraux pour conduire à la spécification du (ou des) plans d'analyse d'un protocole.

Dans le cas général, l'expérimentaliste décrit l'ensemble de ses données à partir d'un certain nombre de facteurs élémentaires, dont la définition est plus ou moins directement issue du plan d'expérience (plan de recueil de données).

place à aucune marge d'appréciation, précisément par suite de son caractère artificiel.

Du point de vue pratique, l'expérience a montré que ce "problème de frontière" ne soulève pas, en général, de difficultés majeures. Cela dit, les problèmes de frontière sont parfois difficiles à "vivre", et certains chercheurs pourraient souhaiter la mise en place de "garde-fous" susceptibles de servir dans les cas douteux. Cependant, on peut se demander si une telle mise en place, qui risquerait d'être assez lourde, ne conduirait pas, une fois encore, à détourner l'attention sur des points somme toute secondaires, au détriment d'autres points incontestablement plus importants, comme celui (que nous aborderons au chapitre suivant) du caractère limité des conclusions permises par un résultat "significatif" ou "non-significatif".

### Structures de plans privilégiées pour la mise en oeuvre des méthodes d'analyse locale.

La mise en oeuvre des procédures d'analyse locale, sur les données que nous avons traitées plus haut, a été grandement facilitée parce que les analyses ont porté sur des plans dont la structure était privilégiée à un double titre :

- du point de vue ensembliste, il s'agissait de plans quasi-complets (et même complets, pour le protocole de groupe fondamental).

- du point de vue statistique, il s'agissait de la structure  $S\langle G \rangle * T$  ( $S * T$  pour les analyses de groupe,  $S\langle G \rangle$  pour les analyses individuelles).

[En fait, le plan de base décrivant les données d'ensemble :  $R\langle S * A * B * C \rangle$ , était déjà lui-même un plan quasi-complet, et on aurait donc pu procéder aux analyses descriptives, du moins à certaines d'entre elles, d'emblée à partir de ce plan. Mais ce plan n'était pas muni de la structure statistique  $S\langle G \rangle * T$  (car à plus d'un facteur de groupe). C'est pourquoi, pour les analyses inférentielles, nous avons envisagé des protocoles dérivés décrits par des plans appartenant à cette structure].

A ce point de l'exposé, il sera sans doute utile de résumer les principaux avantages de ces structures privilégiées.

La structure de plan quasi-complet conduit au langage des formules,

Le facteur composé de ces facteurs élémentaires est un plan (au sens technique que nous avons donné à ce terme, c'est-à-dire qu'il fournit une description injective des données) ; ce plan, que nous appellerons le plan de base, sera le plus "signifiant" de tous les plans du protocole considéré par la suite. Mais en général, ce plan ne sera pas quasi-complet, pas plus qu'il ne sera muni de la structure  $S\langle G \rangle * T$ .

On pourra alors se fixer comme objectif de parvenir, en vue de l'analyse des données, à partir du plan de base, à un plan, ou à des plans, possédant les structures privilégiées, ou au moins l'une de ces structures (par exemple, pour les premiers examens à vue : au moins la structure de plan quasi-complet) : plans que nous appellerons des plans d'analyse.

Tout d'abord, on pourra se demander si le protocole d'ensemble des données ne peut pas lui-même être décrit par un plan ayant la ou les structures privilégiées. Si oui, on prendra ce plan pour plan d'analyse. Si cela n'apparaît pas possible, on pourra faire entrer en considération les principaux effets que l'on souhaite examiner et on recherchera un, ou plusieurs protocoles dérivés susceptibles d'être décrits par un plan privilégié.

Etant donné la double propriété que l'on peut requérir d'un plan d'analyse, on voit que la recherche précédente pourra être effectuée à deux niveaux.

Dans l'exemple des "données H & B", nous nous étions placés (pour les besoins de la cause) dans un "bon cas" : le protocole d'ensemble des données peut déjà être décrit par un plan quasi-complet, à savoir  $R\langle S * A * B * C \rangle$  [en l'occurrence, cette description se trouve tout à fait proche du "plan de base", qui outre les 5 facteurs précédents, ferait intervenir essentiellement des facteurs constants (à une seule modalité), certes importants pour l'interprétation, mais superflus pour le traitement ; c'est pourquoi nous avons pu dire, au prix d'une légère extension de langage (c'est-à-dire : "aux facteurs constants près") que le plan précédent était le "plan de base" de l'ensemble des données]. Nous aurions pu prendre ce plan pour procéder à de nombreuses analyses statistiques ; mais du fait que ce plan comportait plus d'un facteur de groupe, nous avons préféré rechercher tout de suite, et nous aidant des questions que nous voulions poser aux données, des protocoles dérivés de la structure  $S\langle G \rangle * T$ . De la sorte, nous avons été conduits à envisager le protocole de groupe fondamental et les protocoles individuels fondamentaux.

Dans le cas général, la recherche du ou des plans d'analyse pourra amener tantôt à écarter certains facteurs (ce qui renvoie, dans le cadre de la formalisation algébrique, à la théorie des facteurs "superflus" et des "plans minimaux", cf. Réf. 1977a), tantôt à dériver sur d'autres facteurs, voire à introduire également de nouveaux facteurs "auxiliaires" (comme par exemple pour les "données de Cochran & Cox", au chapitre VIII).

Nous n'avons pas, pour l'instant, cherché à systématiser dans un cadre général la démarche qui, à partir d'un plan de base et de questions à examiner, conduirait au ou aux plans d'analyse. Gagnerait-on beaucoup à une telle systématisation ? Cette question, pour nous, reste ouverte à l'heure actuelle. L'expérience a montré tout l'intérêt et le profit, parfois inattendu, que pouvait tirer l'expérimentaliste d'une démarche heuristique au cours de laquelle sont souvent progressivement explicités et précisés des objectifs, qui ne gagneraient peut-être pas nécessairement à être spécifiés d'emblée dans le cadre d'une démarche systématique.

## CHAPITRE VI - INTRODUCTION A LA FORMALISATION LINEAIRE ET A LA NOTION GENERALE DE COMPARAISON

### Introduction

La formalisation linéaire constitue la partie centrale de la construction théorique ; elle consiste à proposer, pour des notions introduites dans ce texte dans un cadre souvent particulier et selon une présentation intuitive, (telles que celles d'effet, d'analyse locale, etc.), des définitions formalisées, valables dans un cadre très général.

Une première tentative de formalisation linéaire avait été exposée dans la Réf. 1968 ; une présentation élargie se trouve dans la Réf. 1976b. Dans le présent chapitre, nous nous bornerons à un "aperçu théorique" suivi d'une illustration à propos des "données H & B".

### Aperçu théorique sur la formalisation linéaire : l'analyse des comparaisons comme généralisation de l'analyse de la variance

Dans ce qui suit, nous désignerons par E le support d'un protocole numérique (éventuellement dérivé, et pondéré) : par exemple ("données H & B"), E pourra être  $S \times A \times B \times C$ , ensemble de toutes les combinaisons des fac-

teurs apparaissant dans le plan, ou encore  $A*B*C$ , ensemble des combinaisons des facteurs systématiques, etc.). Nous appelons contraste sur E toute mesure sur E de masse totale égale à 0 ; l'ensemble des mesures sur E, que nous notons  $\mathbb{R}_E$ , est muni d'une structure d'espace vectoriel ; nous appelons alors comparaison sur E tout sous-espace vectoriel de contrastes sur E, et nombre de degrés de liberté de la comparaison, le rang de ce sous-espace. La pondération sur E permet de munir E d'une mesure de référence (ou mesure fondamentale), grâce à laquelle :

- on peut d'une part munir l'espace  $\mathbb{R}_E$  d'une structure euclidienne en définissant un produit scalaire entre mesures (donc en particulier entre contrastes) ; d'où sur  $\mathbb{R}_E$  une notion d'orthogonalité (entre sous-espaces de mesures, donc en particulier entre comparaisons), puis une norme et une métrique ;

- on peut d'autre part associer à chaque mesure sur E sa densité par rapport à la mesure fondamentale ; on considèrera cette densité comme un vecteur de l'espace vectoriel  $\mathbb{R}^E$  dual de  $\mathbb{R}_E$  ; l'application qui à chaque mesure associe sa densité apparaît alors comme l'isomorphisme de  $\mathbb{R}_E$  vers  $\mathbb{R}^E$  induit par le produit scalaire précédent.

Les notions précédentes ne font intervenir que le plan du protocole. Pour faire entrer en ligne de compte les données, on identifiera l'espace des protocoles numériques avec l'espace  $\mathbb{R}^E$  dual de l'espace  $\mathbb{R}_E$  ; inversement, une mesure sur E pourra être considérée comme une forme linéaire sur  $\mathbb{R}^E$ , espace des protocoles sur E.

Du fait que tout contraste sur E, vecteur de  $\mathbb{R}_E$ , peut être représenté dans  $\mathbb{R}^E$  par sa densité, il résulte que toute comparaison  $\mathcal{C}$  sur E (sous-espace vectoriel de  $\mathbb{R}_E$ ) peut être représentée par un sous-espace de  $\mathbb{R}^E$ , qu'on notera également  $\mathcal{C}$ . Le protocole, en tant que vecteur de  $\mathbb{R}^E$ , peut être projeté orthogonalement sur  $\mathcal{C}$  ; on appellera alors somme des carrés, ou inertie associée à la comparaison (pour ce protocole) le carré de la norme de cette projection. On voit ainsi comment la notion de "comparaison" que nous venons de définir constitue une extension de celle de "source de variation" qui apparaît dans l'analyse de la variance (1), et partant, comment l'analyse des comparaisons constituera une extension de l'analyse de la variance.

### Illustration de la formalisation linéaire à propos des "données H & B"

Concrètement, une comparaison correspondra à une demande d'analyse locale d'un effet. A titre d'exemple, nous prendrons, pour les "données H & B", comme support E, un facteur en correspondance bijective avec le facteur conditions  $J=A*B*C$  et plus précisément  $J_{12} = A_2 * B_2 * C_3$ .

Comme protocole auquel on appliquera la formalisation, on pourra envisager aussi bien :

- le protocole (de groupe)  $A_2 * B_2 * C_3 / S$ , dont les 12 observations sont

---

(1) Quant au terme de "source de variation", nous le réutilisons en analyse des comparaisons, mais en le spécialisant différemment.

les valeurs moyennes (moyennage sur les sujets), relatives aux 12 conditions ;

- l'un des protocoles individuels, tels que  $A_2 * B_2 * C_3 / s_1$ , dont les 12 observations sont les valeurs moyennes (moyennage sur les répétitions) pour le sujet  $s_1$ , relatives aux 12 conditions.

Dans ce qui suit,  $A_2 * B_2 * C_3$  désignera donc l'un quelconque de ces protocoles envisageables). On supposera d'abord que ces protocoles sont regardés comme non-pondérés (ou plus généralement, regardés comme munis d'une pondération uniforme).

Dans le tableau V, les colonnes représentent les 12 conditions, énumérées lexicographiquement comme dans le tableau I ; les lignes du tableau représentent divers contrastes sur lesquels nous allons faire quelques commentaires.

Le contraste n° 1 "oppose" les modalités  $b_1$  et  $b_2$  pour la session  $c_1$  ; il engendre la comparaison à 1 d.l. correspondant à l'effet  $B/c_1$  ("B pour  $c_1$ "), équivalent à l'effet  $B(c_1)$  ("B dans  $c_1$ "). Cette comparaison sera également engendrée par tous les contrastes proportionnels au contraste n° 1, en particulier par le contraste n° 1', lequel a la propriété particulière d'être "normalisé" (la somme de ses coefficients positifs est égale à 1) ; appliqué au protocole  $A_2 * B_2 * C_3$ , le contraste n° 1 aura une valeur égale à celle que donnerait le contraste de coefficients (1, -1) appliqué au protocole dérivé (par moyennage)  $B_2/A * c_1$  (on pourra le vérifier numériquement, par exemple sur le protocole  $A_2 * B_2 * C_3 / S$ ).

De même, le contraste n° 2 opposera les modalités  $b_1$  et  $b_2$ , mais cette fois pour la session  $c_2$ , et engendrera la comparaison  $B/c_2$  (ou  $B(c_2)$ ).

Le contraste n° 3 est obtenu à partir des contrastes n° 1 et n° 2 en faisant leur somme (les contrastes étant des vecteurs, il s'agit donc ici d'une somme de vecteurs) ; ce contraste opposera  $b_1$  et  $b_2$  pour l'ensemble des sessions  $c_1$  et  $c_2$ , et engendra la comparaison (elle aussi à 1 d.l.), notée  $B/c_1 c_2$ , correspondant à l'effet de B "pour" (ou : "moyenné sur") l'ensemble des sessions  $c_1$  et  $c_2$ . La même comparaison est engendrée par le contraste normalisé n° 3', lequel, appliqué au protocole, aura pour valeur la valeur que donnerait le contraste (1, -1) appliqué au protocole dérivé (par moyennage)  $B_2/A * c_1 c_2$ .

Le contraste n° 4 oppose de même  $a_1$  et  $a_2$  pour l'ensemble des sessions  $c_1$  et  $c_2$  et engendre la comparaison notée  $A/c_1 c_2$ .

Le contraste n° 5 est obtenu en multipliant terme à terme les contrastes n° 3 et n° 4 ; il engendrera ici la comparaison d'interaction entre les facteurs A et B, pour l'ensemble des sessions  $c_1$  et  $c_2$ , notée  $A.B/c_1 c_2$ . Cette même comparaison est engendrée par tous les contrastes proportionnels au contraste n° 5, en particulier par le contraste n° 5', lequel, appliqué au protocole, donnera la différence des différences entre les observations relatives aux 4 modalités  $a_1b_1, a_2b_1, a_1b_2, a_2b_2$ , après moyennage sur les 2 sessions  $c_1$  et  $c_2$ , c'est-à-dire la valeur que donnerait le contraste  $(1, -1, -1, 1)$  appliqué au protocole dérivé  $A_2 * B_2 / c_1 c_2$ .

Tous les exemples précédents illustrent la notion de comparaison à 1 d.l. Une comparaison à plusieurs d.l. sera définie comme le sous-espace vectoriel engendré par plusieurs contrastes linéairement indépendants, c'est-à-dire comme une somme de comparaisons (comme une comparaison est un sous-espace vectoriel, et non pas un vecteur, une somme de comparaisons est une somme de sous-espaces vectoriels, et non pas une somme de vecteurs).

Ainsi, les contrastes n° 1 et n° 2, qui sont linéairement indépendants (et même orthogonaux) engendrent la comparaison à 2 d.l., notée  $B(c_1 c_2)$ , correspondant à l'effet de B à l'intérieur des sessions  $c_1$  et  $c_2$  ; cette comparaison sera appelée la somme (orthogonale) des 2 comparaisons  $B(c_1)$  ("B à l'intérieur de  $c_1$ ") et  $B(c_2)$  ("B à l'intérieur de  $c_2$ ") ; et on aura l'égalité suivante (où le symbole "+" est utilisé pour désigner une somme de comparaisons orthogonales) entre comparaisons :  $B(c_1 c_2) = B(c_1) + B(c_2)$ . Cette égalité entre comparaisons entraînera des égalités correspondantes, d'une part pour les d.l., d'autre part pour les inerties.

On remarquera que la comparaison  $B(c_1 c_2)$  ("B à l'intérieur de  $c_1$  et  $c_2$ "), à 2 d.l., est différente de la comparaison  $B/c_1 c_2$  ("B pour  $c_1$  et  $c_2$ ") laquelle est à 1 seul d.l. (on voit pourquoi l'équivalence entre "/" et "()" ne vaut que dans le cas où, à la droite du symbole "/" ou à l'intérieur de la parenthèse, on a une seule modalité).

Cas d'un protocole pondéré ; procédure de remontée d'un contraste.

Les contrastes que nous avons écrits ci-dessus correspondront à des demandes d'analyse "équi pondérée", ce qui rend leur interprétation intuitive.

Pour donner une idée des procédures applicables à un protocole pondéré (avec une pondération non-uniforme), nous reprendrons le protocole individuel  $A_2 * B_2 * C_3 / s_1$ , en supposant maintenant que l'on prend les effectifs comme pondération.

Pour les illustrations qui suivent, nous aurons besoin seulement des 8 effectifs relatifs aux 8 dernières conditions que nous reproduisons ici à partir du tableau I :

a1b1c1	a2b1c1	a1b2c1	a2b2c1	a1b1c2	a2b1c2	a1b2c2	a2b2c2
71	24	70	23	72	24	72	23

Pour construire les contrastes 1 à 4, on pourra utiliser une méthode qui prolonge directement celle suivie dans le cas d'une pondération uniforme.

Ainsi, pour opposer les modalités b1 et b2, pour la sessions c1, c'est-à-dire pour engendrer la comparaison (pondérée) B/c1, on devra maintenant, au lieu de l'un des contrastes n° 1 ou n° 1' écrits plus haut, prendre le contraste n° 1 (P) ("P" pour "pondéré") (ou, bien entendu, tout contraste proportionnel) ; de même que le contraste n° 1', le contraste n° 1 (P) est normalisé, car :  $95 = 71 + 24$  et  $93 = 70 + 23$ .

On obtiendra de façon analogue :

- le contraste normalisé n° 2 (P) qui oppose b1 et b2 pour la session c2, et engendre la comparaison pondérée B/c2 (on aura ici :  $96 = 72 + 24$ , etc.).

- le contraste normalisé n° 3 (P) qui oppose b1 et b2 pour les sessions c1 et c2, et engendre la comparaison pondérée B/c1 c2. (On aura ici :  $191 = 71 + 24 + 72 + 24$ , etc.).

- le contraste normalisé n° 4 (P) qui engendre A/c1 c2.

Mais quand on en viendra à la recherche d'un contraste d'interaction : A.B/c1 c2, il ne sera plus possible d'utiliser la méthode du produit terme à terme (En effet, ce produit n'est même plus, en général, un contraste, et effectivement, dans le cas présent, on peut vérifier que la somme des produits

terme à terme des contrastes A/c1 c2 & B/c1 c2 n'est pas nulle).

La recours à des procédures plus générales s'avère alors indispensable. Ci-dessous, nous allons illustrer, à propos de la recherche d'un contraste engendrant A.B/c1 c2, la procédure générale de remontée d'un contraste (procédure qui, bien entendu, aurait pu être utilisée pour obtenir tous les contrastes écrits précédemment). L'idée générale en est la suivante : on écrit le contraste sur un support (dérivé) pour lequel on dispose d'une définition privilégiée, puis, à l'aide de la mesure-effectifs, on construit un contraste équivalent sur le support primitif.

Dans le cas présent, le support dérivé sera  $A_2 * B_2 / c1 c2$  (en correspondance bijective avec  $A_2 * B_2$ ), et le support primitif  $A_2 * B_2 * C_3$ .

- sur le support dérivé, on dispose d'un contraste privilégié (donc la définition est intrinsèque, en ce sens qu'elle ne fait pas intervenir les effectifs) ; on écrira les coefficients de ce contraste :

a1b1	a2b1	a1b2	a2b2
1	-1	-1	1

- on écrit ensuite les effectifs relatifs au support dérivé :

143	48	142	46
-----	----	-----	----

- on calcule ensuite la "densité du contraste" (en divisant chaque coefficient par l'effectif correspondant, toujours sur le support dérivé):

$\frac{1}{143}$	$\frac{-1}{48}$	$\frac{-1}{142}$	$\frac{1}{46}$
-----------------	-----------------	------------------	----------------

- on "remonte la densité" sur la partie du support primitif correspondant au support dérivé :

a1b1c1	a2b1c1	a1b2c1	a2b2c1	a1b1c2	a2b1c2	a1b2c2	a2b2c2
--------	--------	--------	--------	--------	--------	--------	--------

$\frac{1}{143}$	$\frac{-1}{48}$	$\frac{-1}{142}$	$\frac{1}{46}$	$\frac{1}{143}$	$\frac{-1}{48}$	$\frac{-1}{142}$	$\frac{1}{46}$
-----------------	-----------------	------------------	----------------	-----------------	-----------------	------------------	----------------

- on remultiplie par les effectifs du support primitif (de manière à obtenir à nouveau un contraste) :

$\frac{71}{143}$	$\frac{-24}{48}$	$\frac{-70}{142}$	$\frac{23}{46}$	$\frac{72}{143}$	$\frac{-24}{48}$	$\frac{-72}{142}$	$\frac{23}{46}$
------------------	------------------	-------------------	-----------------	------------------	------------------	-------------------	-----------------

- Enfin, on complète par des zéros sur la partie du support primitif qui ne porte pas la comparaison recherchée (en l'occurrence, les 4 conditions a1b1c0, a1b2c0, a1b2c0, a2b2c0) d'où le contraste n° 5 (P)

reproduit dans le tableau V.

La procédure de remontée nous permettra de souligner une propriété importante de la formalisation linéaire. La supériorité de la procédure de remontée, pour définir l'interaction, par rapport à la procédure "intuitivement envisageable" du produit terme à terme, provient du fait qu'elle met en oeuvre la structure de dualité dans les espaces linéaires (ce qui s'exprime, au niveau de la procédure, par le calcul de la densité).

Manifestement, tant qu'on reste à des situations élémentaires (plans équilibrés, ou orthogonaux, etc.), le recours à cette dualité, pour fournir des critères de choix permettant de retenir les "bonnes notions" (et partant les "bonnes procédures"), n'apparaîtra pas d'une évidente nécessité, car les "bonnes notions" sont alors en quelque sorte "surdéterminées" et on peut, pour les exprimer, se borner aux seules ressources du calcul matriciel, ce que font les théories traditionnelles. Mais il n'en va plus de même pour des situations plus complexes, et c'est alors qu'apparaîtra la pleine portée de la formalisation linéaire.

#### Langage des comparaisons (sur un plan quasi-complet).

Replacées dans le cadre général de la formalisation linéaire, les dérivations élémentaires (restriction et moyennage) apparaissent comme des dérivations linéaires particulières, associées à des comparaisons liées à la structure des facteurs du plan.

La notion générale de comparaison permettra d'enrichir le langage élémentaire des plans quasi-complets (avec les symboles "<>", "+", et "/"), en y adjoignant notamment :

- l'interaction entre 2 facteurs croisés (à un nombre de modalités quelconque) qu'on représentera par le point "." ;
- la somme (orthogonale) des comparaisons, représentée par "+" ;
- la notion de comparaison résiduelle d'une comparaison, représentée par "-" ;
- la notion générale de comparaison intra, représentée par les parenthèses "()" ;

La "brochure verte" (Réf. 1976a) pourra servir de "document de base" pour l'usage du programme VAR3 ; ce texte sera complété utilement par des textes exposant des exemples d'application des programmes à des données réelles - Cf. notamment J-M Hoc, 1975 et B. et M-P. Lecoutre, 1977.

## CHAPITRE VII - PROLONGEMENTS RECENTS ET EN COURS : ANALYSE DESCRIPTIVE DES COMPARAISONS ET ANALYSE FIDUCIAIRE

Les prolongements récents et en cours, tant théoriques qu'informatiques, de l'analyse des comparaisons, se poursuivent principalement dans deux directions, dont le point de départ commun a été une remise en cause des tests de signification en tant que critère d'évaluation des effets. Il nous est en effet apparu de plus en plus clairement que, par rapport aux objectifs de nombreuses recherches expérimentales, le critère des tests de signification pouvait être parfois, dans un certain sens, "trop exigeant", et souvent, dans un autre sens, "trop peu exigeant".

Le critère des tests est parfois "trop exigeant", parce qu'il nécessite un modèle d'échantillonnage, et que dans certaines situations, la validité de ce modèle peut être sujette à caution (\*) ; et pourtant, dans ces situations, des procédures simplement descriptives pourraient souvent apporter déjà des résultats dignes d'intérêt.

Mais le critère des tests est également "trop peu exigeant" lorsque, situation fréquente, la question essentielle qui se pose, à propos de l'effet examiné, est celle de l'importance de cet effet, et non pas de sa seule existence. Or, le test de signification répond à cette seule dernière question (même si c'est au niveau de la population) ; nous y reviendrons plus loin.

### Analyse descriptive des comparaisons

La recherche d'un critère d'évaluation visant simplement à cerner l'importance descriptive d'un effet nous a conduits à développer l'analyse descriptive des comparaisons.

---

(\*) même si, comme nous l'avons rappelé au chapitre 1er, les situations expérimentales dans leur ensemble apparaissent à cet égard comme relativement privilégiées.